Cerca i Anàlisi d'Informació (CAI)
Grau en Ciència i Enginyeria de Dades, UPC

Session 8: Network Analysis

Exercise List, Fall 2019

Basic comprehension questions.
Check that you can answer them before proceeding

1. Make sure you know how to generate Erdős-Rényi, Watts-Strogatz, and Barabasi-Albert networks.
2. Enumerate basic differences between "real-world" networks and Erdős-Rényi networks.
3. Define what a *small-world* network is.
4. Define what a *scale-free* network is.
5. Define the three notions of centrality seen in class

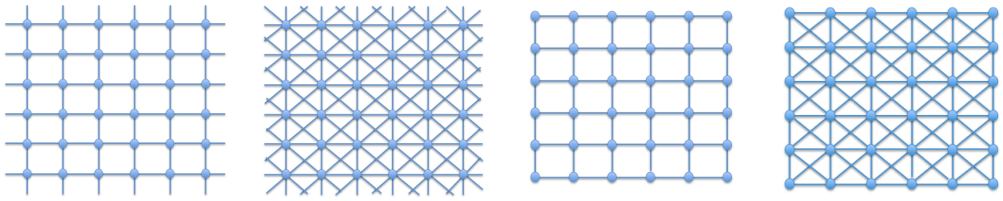
Exercise 1

Tell which of the networks in the following table are random according to the ER model seen in class. In the table, n stands for the number of nodes, m for the number of edges, and C for the clustering coefficient. Justify your answer.

	n	m	C
X1	4941	6594	0.08
X2	125	560	0.07
X3	256985	7778954	0.009

Exercise 2

Compute the global and local clustering coefficient of the following networks:



(1) (2) (3) (4)

1. Infinite two-dimensional grid, where each node is connected to its 4 closest neighbors
2. Infinite two-dimensional grid, where each node is connected to its 8 closest neighbors
3. $n \times n$ two-dimentional grid, where each internal node is connected to its 4 closest neighbors, nodes along the sides are connected to 3 closest neighbors, and nodes in the corners have 2 neighbors.
4. $n \times n$ two-dimentional grid, where each internal node is connected to its 8 closest neighbors, nodes along the sides are connected to 5 closest neighbors, and nodes in the corners have 3 neighbors.

Exercise 3

Design an algorithmic test that, given an undirected network represented by its adjacency matrix, determines whether or not the input network is likely to be a real-world network. What is the complexity of your algorithm?

(Note: There is more than one reasonable answer.)

Exercise 4

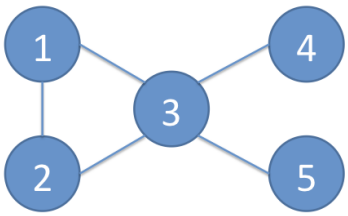
In the following scenarios, which of the centrality measures (degree, betweenness, closeness, PageRank) would you use and how. Please justify your answer.

1. You are a Government Agency, and you want to disrupt a peer to peer network.
2. You are an extremely curious person, and you want to know all the gossip around your network of friends.
3. You are a good doctor, and you want to stop an epidemic (nodes in the network are people, edges are physical connections between people).

4. You like to be in the spotlight, and in order to do that you need to be the only one spreading all the gossip around your network of friends.

Exercise 5

Use your favourite community structure finding algorithm to find 2 communities in the following simple network. Describe in detail what you have done.



Exercise 6

The definition of betweenness is not very robust because we only look at the shortest paths. But a node should be considered central if it is in paths that are not the shortest, but the shortest +1, or shortest +2, etc.

Give a formal definition (or more) of extended centrality that is more robust in this way, by taking into account several paths and not just the shortest one.

Exercise 7

At the end of the slides we briefly defined two problems: selecting a set that minimizes the reaction time to outbreaks, and selecting a set that maximizes the spread of information. Unfortunately, these definitions were not formal.

The bad news is that all reasonable definitions you come up with lead to NP-complete problems, so we will have to rely on approximate or heuristic algorithms.

Give heuristics for these loosely defined problems that you think might work well in practice. Hint: Choosing nodes with a lot of centrality work well. Greedy should help.

(Note: There may be many different decent answers).